NCT number: **NCT02611687**

Protocol ID: **P11-06/BF2.649**

Title: **Double blind, multicentre, randomized, placebo-controlled trial to evaluate safety and efficacy of pitolisant in children from 6 to less than 18 years with narcolepsy with/without cataplexy, followed by a prolonged open-label period.**

Document: **Statistical Analysis Plan, Version 2.0_15 November 2021**

KEYRUS
life science

STATISTICAL ANALYSIS PLAN

BIOPROJET - 1106
PROTOCOL P11-06/ BF2.649
EUDRACT NUMBER: 2013-001506-29

Double blind, multicenter, randomized, placebo-controlled trial to evaluate safety and efficacy of pitolisant in children from 6 to less than 18 years with narcolepsy with/without cataplexy, followed by a prolonged open-label period.

Final version v2.0
November 15th, 2021

Sponsor:
Bioprojet Pharma
9 Rue Rameau
75002 Paris
France

Author(s):
[redacted]

APPROVAL FORM

	Date	Signature
Keyrus Life Science		
[redacted]	15/11/2021	[redacted]
Bioprojet		
[redacted]	15/11/2021	[redacted]
[redacted]	15/11/2021	[redacted]

TABLE OF CONTENTS

1 List of abbreviations and definition of terms 5

2 Introduction 6

3 Study description 6

3.1 Study objectives 7
3.1.1 Primary objective 7
3.1.2 Secondary objectives 7

3.2 Study design 7

3.3 Study plan 8
3.3.1 Study duration 8
3.3.2 Data collected 8
3.3.3 Administration of IMPs 11

3.4 Changes in the conduct of the study 11

3.5 [redacted] 11

4 Statistical methods 11

4.1 General statistical considerations 11
4.1.1 Software used 11
4.1.2 Descriptive statistics 11
4.1.3 IMP groups comparison 12
4.1.3.1 Main Inferential model 12
4.1.3.2 Multiple Testing 12
4.1.4 Missing data 13
4.1.4.1 General considerations 13
4.1.4.2 Dates 14
4.1.5 Derived variables 15
4.1.5.1 General definitions 15
4.1.5.2 First and last IMP intake dates 16
4.1.5.3 Completers 16
4.1.5.4 Baseline characteristics 16
4.1.5.5 Efficacy data 18
4.1.5.6 Safety data 20
4.1.6 Miscellaneous 22

4.2 Sample size calculation 22

4.3 Analysis populations and subgroups 22

4.4 Protocol deviations 23

4.5 Statistical analyses 23
4.5.1 Disposition and Baseline characteristics 23
4.5.1.1 Patients disposition 24
4.5.1.2 Protocol deviations 24
4.5.1.3 Demographic data and other baseline characteristics 24
4.5.2 Treatments of patients 24
4.5.2.1 Extent of exposure and treatment compliance 24

4.5.2.2 Concomitant treatments 24
4.5.3 Efficacy analysis 24
4.5.3.1 Primary endpoints 24
4.5.3.1.1 Primary analysis 25
4.5.3.1.2 Sensitivity and other analyses 25
4.5.3.2 Secondary endpoints 26
4.5.3.2.1 Pediatric Daytime Sleepiness Scale scores (PDSS) 26
4.5.3.2.2 Ullanlinna Narcolepsy Scale Cataplexy subscore (UNS-CTP) 26
4.5.3.2.3 Weekly rate of Cataplexy (WRC) 27
4.5.3.2.4 Maintenance of Wakefulness Test 30-minute version 27
4.5.3.3 Exploratory analysis 28
4.5.3.3.1 Ullanlinna Narcolepsy Scale EDS subscore (UNS-EDS) 28
4.5.3.3.2 Clinical Global Impression of improvement 29
4.5.3.3.3 Response to treatment according to the CGI-C 29
4.5.3.3.4 Child and Adolescent Sleepiness Scale (CASS) 29
4.5.3.3.5 Test of Everyday Attention for Children (TEA-Ch) 30
4.5.3.3.6 Patient sleep diary 30
4.5.3.3.7 Patient's Global Opinion on the effect of treatment 30
4.5.3.3.8 Withdrawal symptoms questionnaire (DSM IV) 31
4.5.4 Safety data 31
4.5.4.1 Adverse events 31
4.5.4.2 Vital signs 32
4.5.4.3 Laboratory evaluations 32
4.5.4.4 Physical examination 33
4.5.4.5 Electrocardiogram 33
4.5.4.6 Childhood depression inventory (CDI) 34
4.5.4.7 Columbia-suicide severity rating scale (C-SSR) 34
4.6 Modifications from the statistical section of the protocol 34
4.7 Interim analysis 34
5 Tables, Listings and Graphs (TLG) 34
6 Template of tables, listings and graphs 34
7 References 34

1 List of abbreviations and definition of terms

Abbreviations	Definitions
AEs	Adverse events
ANCOVA	Analysis of Covariance test
BMI	Body mass index
CASS	Child and Adolescent Sleepiness Scale
CDI	Childhood Depression Inventory
CGI-C	Clinical Global Impression of Change
CGI-S	Clinical Global Impression of Severity
CNS	Central nervous system
CRF	Case report form
CRO	Contract Research Organization
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
DSM-IV	Diagnostic System Medical, fourth version
DSMB	Data Safety Monitoring Board
ECG	Electrocardiogram
EDS	Excessive Daytime Sleepiness
FAS	Full Analysis Set
GCP	Good Clinical Practice
ICH	International Conference on Harmonisation
ICSD	International Classification of Sleep Disorders
IEC	Independent Ethics Committee
IMP	Investigational Medicinal Product
KLS	Keyrus Life Science
MAR	Missing At Random
MCMC	Markov chain Monte Carlo
MD	Missing Data
MI	Multiple Imputation
MNAR	Missing Not At Random
MSLT	Multiple Sleep Latency Test
MWT	Maintenance of Wakefulness Test
OSA	Obstructive Sleep Apnoea syndrome
PDSS	Pediatric Daytime Sleepiness Scale
PPS	Per Protocol Set
PS	Paradoxical sleep
PSG	Polysomnography
REM	Rapid Eye Movement
RS	Randomised Set
SAF	Safety Set
SAE	Serious Adverse Event
SD	Standard Deviation
TEA-Ch	Test of Everyday Attention for Children
UNS	Ullanlinna Narcolepsy Scale
UNS-CTP	Ullanlinna Narcolepsy Scale - Cataplexy
UNS-EDS	Ullanlinna Narcolepsy Scale - Excessive Daytime Sleepiness
TLG	Tables, Graphs and Listings
WRC	Weekly rate of Cataplexy

2 Introduction

This document describes the frame of the statistical analyses that will be conducted for the study P11-06/ BF2.649: A double blind, multicenter, randomized, placebo-controlled trial to evaluate safety and efficacy of pitolisant in children from 6 to less than 18 years with narcolepsy with/without cataplexy, followed by a prolonged open-label period.

This Statistical Analysis Plan has been written in agreement with:

- The Clinical Study Protocol Version 3.0 dated September 30th, 2020
- The Case Report Form (CRF) Version 7.0 dated December 16th , 2020

This document is the reference document for all the statistical analyses of **the blind part** of this study.

The purpose of this document is to describe:

- The characteristics of the study as defined in the protocol: objectives, design and conduct of the study
- The criteria/variables analysed and the derived variables to be created.
- The planned statistical analysis and the methodology to be used.

3 Study description

According to guidelines published by European task force *[Billiard M et al, Eur J Neurol. 2006]* [i], management of narcolepsy with or without cataplexy relies on several classes of drugs, namely stimulants for EDS, antidepressants for cataplexy and hypnosedative drugs for disturbed nocturnal sleep *[Dauvilliers Y et al, Clin Neurophysiol. 2003]* [ii].

Narcolepsy is a neurological disorder frequently occurring from childhood and persisting through adolescence and adulthood. Individuals suffering from narcolepsy exhibit excessive daytime somnolence, sleep attacks, cataplexy, dysomnia, metabolic perturbations including weight gain, and problems in social interaction and academic performance. The prevalence of narcolepsy in childhood is not known but can be estimated from adult studies to be greater than 20-60 per 100,000 in Western countries. (Handb Clin Neurol. 2013; 112:839-45. doi: 10.1016/B978-0-444-52910-7.00003-9. Pediatric narcolepsy: clinical and therapeutical approaches. Lecendreux M.)

Narcolepsy occurs during childhood in combination with cataplexy in one-third of the subjects. Symptoms may develop rapidly over a few weeks or months, with excessive daytime sleepiness and cataplexy being the most dramatic and observable symptoms. It can be secondary to brain tumors or several rare diseases, but in most cases narcolepsy with or without cataplexy is a primary condition, better explained by the selective loss of hypocretin neurons in posterolateral hypothalamus. (Paediatr Drugs. 2014 Oct; 16(5):363-72. doi: 10.1007/s40272-014-0083-3. Pharmacological management of narcolepsy and cataplexy in pediatric patients. Lecendreux M)

Pitolisant is an H3R inverse agonist that promotes significantly vigilance in narcoleptic patients. The results obtained in the previous studies showed that BF2.649 reduced significantly the diurnal somnolence compared to placebo confirming its wakening effect against EDS with a gain of 4.5 points on the ESS and demonstrating its anti-cataplectic effect when administrated on an individual titration scheme established on basis of individual benefit/tolerance ratio**.** Phase III studies compared pitolisant to Placebo and to modafinil the current first line treatment of narcolepsy: the results are in favour of pitolisant compared to Placebo and do not show any statistically significant difference compared to modafinil.

The purpose of this study is to check whether Pitolisant (BF2.649) is efficient to decrease EDS and the supportive symptoms of narcolepsy in children from 6 to less than 18 years of age, as well as to reduce the number of cataplexy crises (for patients with cataplexy).

3.1 Study objectives

3.1.1 Primary objective

The primary objective of this study is to compare pitolisant to placebo, with regard to Excessive Daytime Sleepiness (EDS) on the one hand, and to the number of cataplexy episodes, if any, on the other hand.
This trial will characterize the efficacy of pitolisant (5-, 10-, 20-, 40 mg/d in Double Blind Period and 5-, 10-, 15-, 20-, 30-, 40 mg/d in Open Label Period) compared to placebo in showing an incremental improvement to the situation particularly in terms of a reduction of EDS as measured by the Ullanlinna Narcolepsy Scale (UNS), and also on the number of cataplexy episodes, if any.

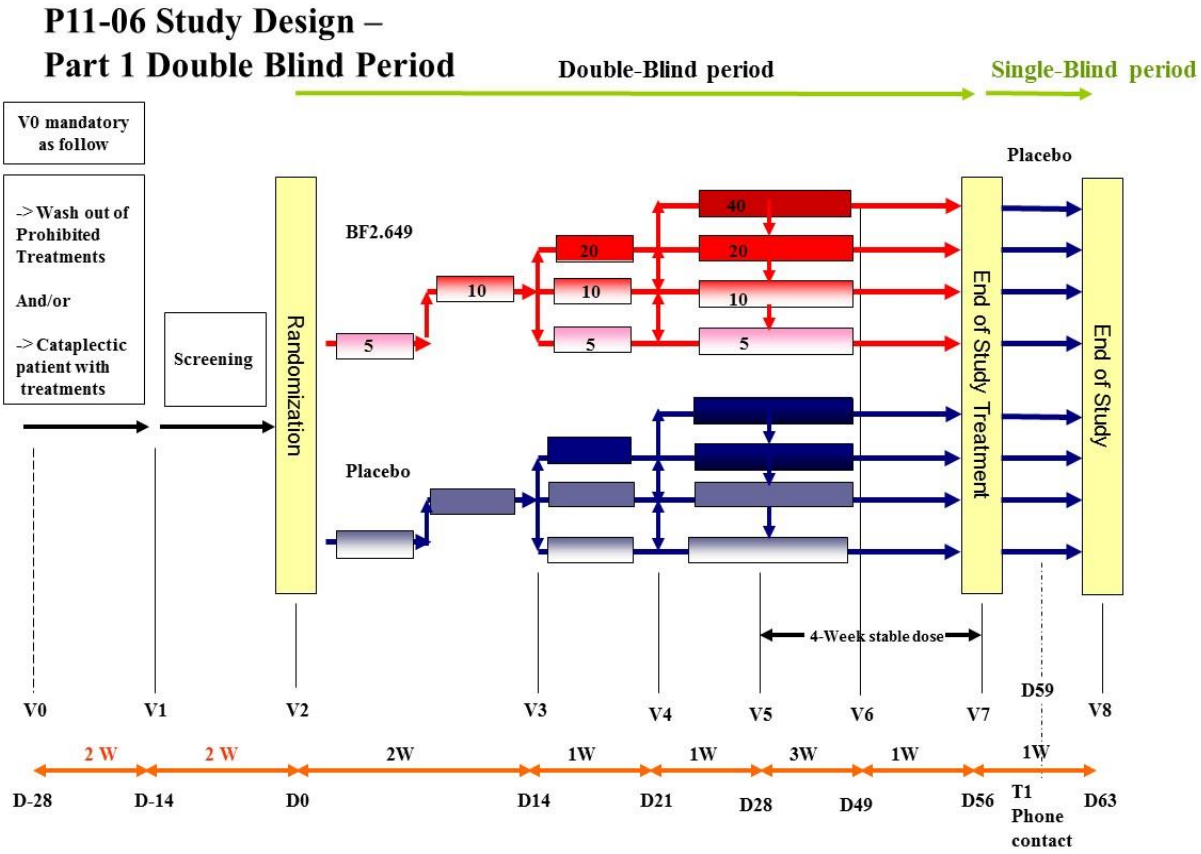
3.1.2 Secondary objectives

The secondary objectives of this study are to assess, between groups:

- Changes in EDS measured:
    - by the maintenance of wakefulness test (MWT)
    - by the Paediatric Daytime Sleepiness Scale (PDSS)
    - by the Child and Adolescent Sleepiness Scale (CASS)
- Changes in number of cataplexy episodes per weeks, recorded in sleep diary by patient and/or parent/teacher;
- Severity of EDS and cataplexy measured by the clinical Global Impression of severity and change
- Differences in weekly frequency of cataplexy episodes, recorded in sleep diary by patient and/or parent/teacher;
- Changes in the TEA-Ch test
- Comparison in withdrawal symptoms questionnaire
- Comparison in patients' Global Opinion on treatment effect at the end of treatment.
- To check the clinical and biological tolerance of pitolisant.

3.2 Study design

This is a double blind, multicenter, randomized trial followed by a prolonged open-label period.
This SAP only deals with the blind part.



3.3 Study plan

3.3.1 Study duration

The expected duration of patient participation from the screening visit (V0) until the final blind visit (V8) will be 13 weeks, including:

- Baseline period without investigational treatments (from V0 to V2): 4 weeks,
- Low and medium dose titration phase (from V2, randomisation, to V3): 2 weeks,
- Individual dose adjustments (from V3 to V5): 2 weeks,
- Stable dose phase (from V5 to V7): 4 weeks,
- 1-week study drug wash out (placebo) after study drug discontinuation.

During the blind phase, the patient will be followed in an ambulatory way by 9 visits (8 visits during double blind period and 1 visit for the single blind period) on the site.

3.3.2 Data collected

A detailed list of procedures performed at each visit is presented in table1.

Table 1: Study Schedule measurement in blind period

Visits / Procedures	Screening V0[8]	V1	Inclusion V2	V3	V4	V5	V6	Endpoint V7	T1- Phone Contact	V8	Premature withdrawal [3]
Study day	D-28	D-14	D0 ±2	D14 ±2	D21 ±2	D28 ±2	D49 ±2	D56 ±2	D59 ±1	D63 ±2	+ 5
Informed consent	X	X*									
Narcolepsy history, Medical history and prior medications	X	X									
Physical examination, vital signs	X	X	X	X	X	X	X	X		X	X
ECG [6]		X	X	X	X	X	X	X		X	X
Lab tests [2]		X						X			X
Selection criteria	X	X	X								
Polysomnography - MSLT			X[7]								
MWT			X					X			
PDSS		X	X	X	X	X	X	X		X	X
CASS		X	X	X	X	X	X	X		X	X
Ullanlinna narcolepsy scale		X	X				X	X			X
TEA-Ch		X					X				
CGI EDS + CGI Cataplexy		X	X				X	X		X	X
Childhood Depression Inventory (CDI)		X			X		X			X	X
C-SSRS		X	X	X	X	X	X	X		X	X
Adverse events and concomitant medications	X	X	X	X	X	X	X	X	X	X	X
Dispensation of study drugs[1]			X	X	X	X		X			
Drug accountability				X	X	X	X	X		X	X
Withdrawal symptoms questionnaire (DSM IV)									X	X	X
Patients' global opinion[5]								X			X
Sleep diary delivery	X	X	X	X	X	X	X	X			
Sleep diary retrieval[4]		X	X	X	X	X	X	X		X	X

1 The 4-week escalating dosage phase is followed by a 4-week stable-dose period during which the dose will be 5-, 10-, 20, or 40 mg/d of pitolisant or placebo
2 Complete biological examination including: Red Blood Cells, White Blood Cells (differential count, absolute value), Hemoglobin, haematocrit, Mean corpuscular volume (MCV), platelet count, quick time, sodium, potassium, chloride, creatinine, alkaline phosphatases, urea, alanine aminotransferase (ALAT), aspartate aminotransferase (ASAT), gamma glutamyl transpeptidase (GGT), total bilirubin, glucose, triglycerides, total cholesterol, βHCG (only for pubescent female). Dipstick urinalysis. A urinary drug testing (test applicable to patients from 12 years).
3 The premature withdrawal from the study should be followed by a visit performed within a maximum of 5 days after the last dose of study drug.
4 The sleep diary has to be completed every evening from V0 to V8.
5 The global opinion will be reported by patient if able to express himself. If not will be reported either by parents or teachers.
6 ECG: calculation of the QTcF value and delta of QTcF values between baseline ECG and ECGs performed at each subsequent visits.
7 Polysomnography-MSLT will be performed the day preceding V2, unless performed during the previous year.
8 V0 is required to ensure proper selection of patients under prohibited medications. For patients not using any prohibited medication study or patients cataplectic not using anti-cataplectic treatments may start at V1.

* Inform consent will be signed once either at V0 or V1 when applicable.

3.3.3 Administration of IMPs

When the patient receives his/her study number, the corresponding investigational treatment identified by the treatment number will be allocated to the patient according to the randomization list established by an independent company and managed by Interactive Web Response Services (**IWRS**) according to randomization.
The study treatment will be dispensed only under the restricted condition defined in the protocol. The treatment will be dispensed only by the investigator under his direct supervision or the pharmacist. Time of dispensation and initials of the person dispensing the drug will be recorded in the CRF. The tear-off label of each pack will be stuck on the prescription form.
At each visit the compliance to the treatment will be investigated. Details of the quantities of each medication dispensed will be entered into the accountability form. The patients will be asked whether the investigational treatment is taken as prescribed. If not, any change in the treatment and the number of forgotten tablets will be recorded in the case report form (CRF). In addition, the number of remaining tablets in the pillbox will be counted and compared to the theoretical number. Any mismatch will be investigated with the patient.

3.4 Changes in the conduct of the study

The last version of the protocol is version 3.0, dated September, 30th 2020. The original protocol was amended 3 times. The list of exhaustive changes is described in the summary of changes from each protocol amendment.

[redacted]

4 Statistical methods

4.1 General statistical considerations

4.1.1 Software used

All statistical analyses will be performed with the SAS® software version 9.4 or higher.

4.1.2 Descriptive statistics

Continuous variables will be described using: number of non-missing observations (N), number of missing observations (Nmiss), arithmetic mean (Mean), standard deviation (SD), minimum (MIN), quartile 1 (Q1), median (Median), quartile 3 (Q3) and maximum (MAX). MIN, Q1, Median, Q3 and MAX will be given at the data number of decimal points. Mean and SD will be given at 1 additional decimal point. 95% two-sided confidence interval (CI) will be calculated when appropriate using the standard method (Standard normal distribution).

Categorical variables will be presented using N, Nmiss and percentages (%). Missing data or unknown responses will not be included in the denominator for the calculation of percentages but will be presented in the results tables. Proportions will be displayed with one decimal.

4.1.3 IMP groups comparison

4.1.3.1 Main Inferential model

Unless otherwise specified, all the endpoints referred in detail in the following section will be analyzed according the same statistical model allowing a comparison of results between the studied endpoints. The significance of the difference between the IMP compared with placebo will be assessed through a general model of analysis of Covariance (ANCOVA) in which the studied endpoint at final time (noted Y_f) constitutes the dependent variable of the model. Independent variables are (1) the baseline value of the studied endpoint (noted Y_b) analyzed as a fixed covariate, (2) the treatment effect (Noted Trt) coded 0 (placebo) or 1 (IMP) (cell mean contrast formulation), and (3) the center considered as a random factor of the intercept constituted by a fixed part k_0 and a random component $N(0,\sigma_c)$. This model is summarized by the following expression:

$$Y_{if} = (k_0+\zeta) + k_b Y_b + k_t Trt + \epsilon_i \quad \epsilon_i = N(0,\sigma) \text{ and } \zeta = N(0,\sigma_c) \quad \text{[general Model -1]}$$

Note 1-This model does not assume a treatment-baseline interaction term (as required in CPMP, 2003). Our decision of using center as a random factor is justified by the useless estimate of each center, in focusing on the measurement of the standard deviation of the intercept. The treatment will be considered as fixed factor without center random component.
Note 2- This model is a Linear Mixed Model (LMM) as it combines random and fixed factors, assumes linearity of baseline and hypothesizes the studied endpoint distributed according to a normal (Gaussian) distribution. This model constitutes the main model used in this trial for all the studied endpoints unless otherwise specified.

4.1.3.2 Multiple Testing

(ICH E3; 9.7.1, 11.4.2.5. ICH E9; 2.2.5, FDA-2017)
The main analysis is defined as the analysis of covariance following general Model 1 [section 4.1.3.1] conducted on the main selection (FAS), on the primary endpoint (UNS) considered as a global endpoint aggregating both EDS and cataplexy symptoms. Secondary endpoints will be PDSS and the UNS-CTP subscale measuring the IMP effect on EDS and cataplexy, respectively. Two other secondary endpoints will be the weekly rate of cataplexy measured by the patient diary (WRC) and the MWT. All the other endpoints will be considered as exploratory and supportive.
The main endpoint (UNS) and the four secondary endpoints necessitate 5 tests. We will use the fixed-sequence strategy testing the endpoints in the pre-defined order UNS, PDSS, UNS-CTP, WRC and MWT, each one tested at two-sided α=.05 significance, moving to the following endpoint conditionally to a significant difference reached on the previous endpoints (FDA, section C.5). This strategy protects the results with type 1 risk bounded by P=.05.
Our choice is justified as follows: UNS constitutes a global appraisal of both EDS and cataplexy in pediatric narcolepsy and must regarded as a relevant general first endpoint. PDSS and UNS-CTP are validated scales to measure EDS alone and cataplexy alone, respectively. Conditionally to a significant effect on the main UNS endpoints, the effect of the IMP will be assessed on the two symptoms, separately. Finally MWT and WRC are alternative endpoints to measure EDS and cataplexy, respectively, and will be assessed conditionally to the significance obtained on UNS, PDSS and UNS-

CTP. Besides, EDS is a more frequent symptom, thus it was tested before cataplexy, and PDSS and UNS-CTP were shown with a higher power than MWT and WRC, respectively [ref-validation]

4.1.4 Missing data

4.1.4.1 General considerations

(ICH E3; 9.7.1, 11.4.2.2. ICH E9; 5.3. EMA Guideline on Missing Data in Confirmatory Clinical Trials)

Missing data will be imputed for five main endpoints of this study (UNS, PDSS, UNS-CTP, WRC and MWT). For the other endpoints no missing value will be imputed.

A Multiple Imputation (MI) procedure will be used to impute the missing evaluations, as a prior step, assuming that those patients would be in their randomized arm. The partially complete datasets will be generated with a model including the following covariates age, gender and the main model components (section 2.2).

A total of 100 imputed partially complete datasets will be generated.

MI inference involves 3 consecutives phases:

- 1/ Imputation step:

The missing pattern is supposed to be monotone, so the regression method will be used to impute missing data.

The partially complete datasets will be generated with a model including age, gender factors and taking into account the baseline score of the primary endpoint, based on patient which every post-baseline measurement available. It is of note that this imputation step might be preceded by one MI approach based on MCMC method, in case of arbitrary missing pattern. A total of 100 imputed partially-complete datasets will be generated.

The mean MD imputation will based on Missing at Random (MAR) corresponding to the hypothesis that missingness is assumed independent of the outcome, or to estimate a “de jure” (per protocol) estimand, censoring after any protocol deviation. This hypothesis was considered as the most likely in this particular context.

Despite the likelihood of this option, a secondary MD imputation will be conducted according to the Missing at Random hypothesis (MNAR), for which several options to construct the joint distribution of the observed and unobserved data have been proposed, each reflecting a different MNAR mechanism (Carpenter et al., 2013). This appropriate choice was discussed in the general context of narcolepsy. We assumed that after dropping out, participants from the active arm have the same outcomes as the reference group individuals, hypothesizing that the treatment effect is lost after the individual leaves the study. This option is called Jump to reference (J2R), after dropout, the participant's conditional outcomes are associated to those of the control arm. The joint distribution is a Multivariate Normal model with mean parameters from the randomised arm until date of interruption (Di) and from the reference group afterward. The covariance matrix corresponds to the parameters from the randomised arm until Di and to the reference group for the conditional components of the post-dropout variables, given the pre-dropout measurements.

Concerning the J2R, the jump-to-placebo approach will be set up for these patients’ imputations, by using the %mistep macro from James Roger (https://www.lshtm.ac.uk/research/centres-projects-groups/missing-data#dia-missing-data , zip file MISTEP20180327).

Usualy, a sequence of %mistep is run, the subsequent models include the residuals of the previous model. In this study only one model will be used (for the endpoint value imputation).

In case of imputed data outside the range, the minimun or the maximun bound will be used.

This macro will be used because no SAS procedure are currently available to perform the jump to reference approach.

- 2/ Analysis step:

The same analysis as described above will be applied to each of the 100 imputed datasets.

- 3/Combination step:

Statistical inferences will be generated by combining results from 100 analyses using Rubin's formulae (Rubin, 1987).
The multiple imputation estimator of the difference between active treatment and placebo is the average of the individual 100 estimators. The variance of the estimator is the combination of the between and within-imputation variability (Kenward & Carpenter, 2007).

4.1.4.2 Dates

- **First and last IMP**

After selection of the dates of first and last IMP intake as defined below, if these dates are missing or incomplete, the following substitution rules will be applied:

Date to substitute		Substituted date
First IMP intake	../mmm/yyyy	Date of randomisation if complete with same month and year *Otherwise:* **01/mmm/yyyy**
	../.../yyyy	Date of randomisation if complete with same year *Otherwise:* **01/JAN/yyyy**
	../.../....	Date of randomisation if complete
Last IMP intake	../mmm/yyyy	**Last available date*** if same month and year *Otherwise:* **last day of the month/mmm/yyyy**
	../.../yyyy	**Last available date*** if same year *Otherwise:* **31/DEC/yyyy**
	../.../....	**Last available date***

Last available date (only for patients randomised) is defined:
date of death if patient died, else last visit date on the V2-V7 period

Notes:

- ../mmm/yyyy = missing day,
  ../.../ yyyy = missing day and month,
  ../.../.... = missing date.

- Missing dates will be substituted only for patients having taken at least one dose of IMP.

- **Previous and concomitant treatments dates**

Substitution rules for previous and concomitant treatments intake dates

Date to substitute		Substituted date
Start date	../mmm/yyyy	**01/mmm/yyyy**
	../.../yyyy	**01/JAN/yyyy**
	../.../....	If stop date is non-missing and inferior to informed consent date then: **Stop date** Else: **Informed consent date**

Stop date	../mmm/yyyy	If patient died same month and year then
		Date of death
		Else
		Last day of the month/mmm/yyyy

	../.../yyyy	If patient died same year then
		Date of death
		Else
		31/DEC/yyyy

	../.../....	If patient died then
		Date of death
		Else
		No substitution
		(*i.e.*, treatment considered as still on-going)

Note: ../mm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = missing date

- **Substitution rules for diagnosis dates**

Date to substitute		Substituted date
Date	../mmm/yyyy	01/mmm/yyyy
	../.../yyyy	01/JAN/yyyy
	../.../....	No substitution

4.1.5 Derived variables

4.1.5.1 General definitions

The following definitions will be considered:

- **Analysable value** will be defined as any non-missing value.
- **Baseline value** will be defined as :
  - For UNS, PDSS and CASS parameters as the mean of V1 and V2 values on condition that values are prior to the first IMP intake (*i.e.* before or the same date as the first IMP intake date). In case of a visit missing value, the baseline will be equal to the non-missing value.
  - For other parameters, the last analysable value prior to the first IMP intake (*i.e.* before or the same date as the first IMP intake date).

  Note: In case of patient randomised but not treated over the blind period: value at baseline is defined as the last analysable value prior or equal to the date of randomisation.
- **Post-baseline value** will be defined as any value recorded at a given timepoint after baseline.
- **Change from baseline** will be defined as the arithmetic difference between a post-baseline value and the baseline value for a given variable at a given time point.
- **End value** will be defined only for UNS, PDSS and CASS parameters as :

- The mean of V6 and V7 values. In case of a visit missing value, the End value will be equal to the non-missing value.

4.1.5.2 First and last IMP intake dates

For patients having taken at least one dose of IMP over the period V2-V7, the dates of first and last IMP intake on the analysis period will be defined as follows:

- The date of the first IMP intake at the first visit performed within the analysis period,
- The date of the last IMP intake at the last visit performed within the analysis period.

After selection of the dates of first and last IMP intake as defined above, if these dates are missing or incomplete, substitution rules will be applied to identify baseline value and emergent adverse events.

4.1.5.3 Completers

- **Completers the double blind period**
  Patients who completed the study until the visit V7.
- **Completers the blind period**
  Patients who completed the study until the visit V8.

4.1.5.4 Baseline characteristics

- **Previous treatments**

The existence of a previous treatment (Yes/No) is defined from the presence, or not, of an Anatomical therapeutic chemical classification and/or Preferred name.
Previous treatment is defined as any treatment with associated stop date strictly inferior to the first IMP intake date.
Note: In case of patient randomised but not treated the previous treatment is defined as any treatment with an associated stop date strictly inferior to the date of the randomisation visit.

- **Concomitant treatment**

The existence of a concomitant treatment (Yes/No) is defined from the presence, or not, of an Anatomical therapeutic chemical classification and/or Preferred name.
The periods considered for the analysis are:

- **At inclusion** for which treatments with start date ≤ inclusion date (V2) and stop date ≥ inclusion date (V2) or missing are taken into account.
- **During the treatment period** for V2-V7 period for which treatments:
  - with start date ≥ first IMP intake date and < last IMP intake date, or
  - with start date ≤ first IMP intake date and stop date ≥ first IMP intake date or missing date

  are taken into account.

Concomitant treatments could be considered in one or several of the possible analysis periods.

Intake of oxybate sodium (Yes/No): A patient will be considered to have taken at inclusion or during the treatment period a concomitant treatment which the Standardized Medication Name="OXYBATE SODIUM".

- **Narcolepsy duration (years)**
  (Inclusion visit date (V2) - confirmed diagnosis date +1) /365.25
- **Study duration (in days)**
  Last visit date until V8– Inclusion visit date (V2) + 1
- **Treatment duration (in days) on V2-V7 period**

Last IMP intake date – First IMP intake date + 1

- **Compliance (%) during the double blind period**

The overall compliance will be computed:

COMP = CD/CT

- **CD** : Number of tablets Dispensed - Unused Tablet number returned, at the end of the double blind period

  With Number of tablets dispensed at each visit according to the dose:

[redacted]

- **CT** : Number of Treatment Exposure Day* Theoric Number of tablets per day, at the end of the double blind period

  Amount predicted according the daily dose:

  - 5 mg: 1 tablet of 5 mg * treatment duration in days at this dose
  - 10 mg: 2 tablets of 5 mg * treatment duration in days at this dose
  - 20 mg: 1 tablet of 20 mg * treatment duration in days at this dose
  - 40 mg: 2 tablets of 20 mg * treatment duration in days at this dose

  Between the V2 and V3 visits there is a dose increase (5 => 10 mg). According to the collected data it's impossible to know when the patient took 1 tablet more. So for the compliance, for the CT computation, the assumption that 1 tablet is taken during the first seven days and 2 after will be done between V2 and V3.

Compliance will be considered as missing for patients with premature withdrawal during the double blind period without information about the number of tablets returned at the premature withdrawal visit.

[redacted]

[redacted]

- **Age (in classes)**
  Two age classes will be computed:
  - [6-11] years,
  - [12-18[years,

- **BMI**

For weight, only one decimal should be kept for the analysed values.
The Body Mass Index is calculated at a visit in the database:
BMI (kg/m²) = Weight (kg) / (Height (cm) x 0.01)²,
using the height and the weight recorded at the corresponding visit.
The result is rounded to one decimal place.
The BMI Classes refer to the statistical methodology used by the WHO standards.
This methodology is computed according to the sex and the age in months.
For each visit, the calculated age at the concerned visit is taken into account.

- If it's a boy, the reference table is the table « BMI-for-age_WHO.xls» sheet « Boys »
- If it's a girl, the reference table is the table « BMI-for-age_WHO.xls» sheet « Girls »

According to these tables:

- If the BMI < -2SD then BMI class= « underweight »
- If the -2SD ≤ BMI ≤ +1 SD then BMI class= « normal range »
- If the +1SD < BMI ≤ +2 SD then BMI class= « overweight »
- If the BMI > +2 SD then BMI class= « obese »

4.1.5.5 Efficacy data

- **Ullanlinna Narcolepsy Scale (UNS)**

For the subscores recalculation the correspondence between the character and numerical modalities:

For questions about « Knees unlocking symptoms » (UNS01), « Mouth opening symptoms » (UNS02), « Head nodding symptoms » (UNS03) and « Falling down symptoms » (UNS04) :

Character	Numeric
NEVER	0
1-5 TIMES DURING LIFETIME	1
MONTHLY	2
WEEKLY	3
DAILY OR ALMOST DAILY	4

For questions about « Fall asleep during reading » (UNS07), « Fall asleep during travelling » (UNS08), « Fall asleep during standing » (UNS09), « Fall asleep during eating » (UNS10) and « Fall asleep during other unusual » (UNS11) :

Character	Numeric
NEVER	0
1-5 TIMES DURING LIFETIMES	1
MONTHLY	2

WEEKLY	3
DAILY OR ALMOST DAILY	4

For the question « Fast to feel asleep the evening » (UNS05):

Character	Numeric
>40 MIN.	0
31-40 MIN.	1
21-30 MIN.	2
10-20 MIN.	3
<10 MIN.	4

For the question « Sleep during the day» (UNS06):

Character	Numeric
NO NEED	0
I WANTED BUT CANNOT SLEEP	1
2 DAYS OR LESS EACH WEEK	2
3-5 DAYS EACH WEEK	3
6-7 DAYS EACH WEEK	4

Two sub-scores are derived:

- ✓ **The cataplexy score**:
  Sum of the first four items (UNS01 + UNS02 + UNS03 + UNS04)
- ✓ **The EDS score**:
  Sum of the seven other items (UNS05 + UNS06 + UNS07 + UNS08 + UNS09 + UNS10 + UNS11)

- **CGI**

Two scores are computed within the ranges of 0 -7 (0 = Not assessed):

- Global Improvement (CGI-I) score
- Severity (CGI-S) score

The response to treatment according to CGI-I is derived at a visit as:

- If CGI-I score = 1 (very much improved) or 2 (much improved) or 3 (minimally improved) then response to treatment at the visit = Yes.
- If CGI-I score > 3 then response to treatment at the visit = No.
- Otherwise, response to treatment at the current visit is not calculated.

- **Pediatric Daytime Sleepiness Scale scores (PDSS)**

The response to treatment according to PDSS is derived at a visit as a score reduction of 3 Between the baseline and the final value:

- If $PDSS_B$-$PDSS_F \geq 3$ then response to treatment at the visit = Yes.
- If $PDSS_B$-$PDSS_F < 3$ then response to treatment at the visit = No.
- Otherwise, response to treatment at the current visit is not calculated.

- **Maintenance of Wakefulness Test 30-minute version**

For this assessment the four test have been considered. Therefore, arithmetic means and sum on these 4 tests are computed.

Mean value for each latency parameter, in minutes, for a visit, will be estimated as the arithmetic mean of the 4 tests.

Before this mean computation each latency parameter (sleep onset, Stade I, Stade II, stade III and SP) must be recalculed to take into account the "Awake" status and the test duration:

- For a test, if "Awake"= "yes" then the sleep onset latency has to be considered at 30min (even if another value is entered).

- If a reported value is greater than 30min this value will be censured at 30min.

Number of Sleep onset in SP for a visit, will be computed as the sum of Yes (=1) of the 4 tests.

A survival response will be computed according to the "sleep onset latency" recalculed:
- Non-response (0=patient remained awake until 30 minutes) = Mean sleep onset latency ≥ 30min.
- Response (1=patient falling asleep before 30 minutes) = Mean sleep onset latency < 30min.

- **Patient sleep diary**

- Rate of cataplexy (WRC), Number of diurnal involuntary sleep attacks and episodes of severe daytime sleepiness (NSAE) and Number of hallucinations and sleep paralysis episodes (NHS) are weekly ratio computed as :
  [Number of episodes of the concerned parameter in the period / Number of days in the period] * 7
  Whith Number of days in the period= Date of the last visit of the period - Date of the first visit of the period + 1
  Results are rounded to to the nearest integer number.

- Night duration (NightD), in hours, is calculated as:
  Wake-time (morning) - Hour of bedtime (evening)
  The mean of night duration per day during the studied period is computed as:
  Sum of night duration in the period / Number of days in the period
  Whith Number of days in the period= Date of the last visit of the period - Date of the first visit of the period + 1
  The result is rounded to one decimal place.
  In case of several hours at the same date the following rule will be applied:
  - For wake-time the last hour will be kept
  - For bedtime the first hour will be kept

4.1.5.6 Safety data

- **<u>Adverse Event</u>**

  - **Treatment Emergent Adverse Event**

A Treatment Emergent Adverse Event (TEAE) corresponds to an Adverse Event (AE) that occurred on or after first dose of treatment up to 7 days post-last dose.

In case of partially/fully missing AE start date:
- Following imputation rules should be applied in order to determine whether the AE is treatment-emergent or not:
  - If only start day is missing, then it will be replaced by the minimum of (last day of the month, last treatment dose intake)
  - If both start day and month are missing then:
    - if year is equal to the year of the first treatment dose intake, then start date will be replaced by the first treatment dose intake

        - if year is greater than the year of the first treatment dose intake and equal to the year of the last treatment dose intake, then start date will be replaced by the last treatment dose intake
        - if year is greater than the year of the first treatment dose intake and smaller than the year of the last treatment dose intake, then start date will be replaced by 1-July
    - If start date if fully missing, then start date will be replaced by the date of the first treatment dose intake
- Emergence will be checked in accordance with the response to the question « Started before study drug intake ». If the response is "Yes" the AE can't be considered as emergent

- **Treatment Emergent Adverse Events related to study treatment**

An event is considered to be related to the study treatment if the assessment of imputability is likely related or possibly related, or if the assessment is missing.

- **Non-Treatment Emergent Adverse Events (non-TEAEs)**

Non-TEAEs are AEs that occurred prior to treatment.

- **Serious Adverse Events**

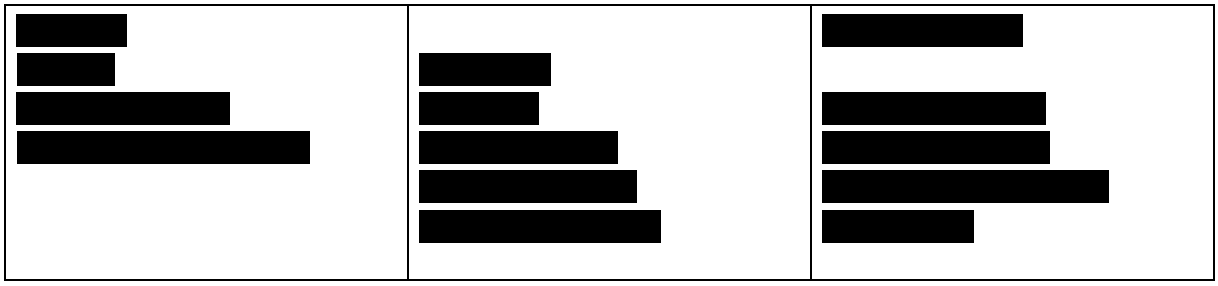
By definition, a serious AE (SAE) is an AE that, at any time, fulfils one or more of the following criteria:

- results in death,
- is life threatening (patient at risk of death at the time of the event),
- requires in-subject hospitalisation or prolongation of existing hospitalisation,
- results in persistent or significant disability/incapacity,
- is a congenital anomaly/birth defect,
- is any medically important.

The seriousness of the event will be directly assessed by the investigator (specific variable within the AE database); therefore no derivation is needed.

- **Special Interest AEs**

Special Interest AEs are AEs with the following MedDRA Preferred Terms (PTs):



- **Treatment Emergent Adverse Events in the stable dose period**

The stable dose is the treatment dose taken during the period V5 to V7.

So TEAE in the stable dose are TEAE occurred between visit V5 (included) and the visit V7 (included)

- **ECG**

The following classes will be derived for **QTCF :** <= 450,] 450; 480],] 480; 500] and > 500

4.1.6 Miscellaneous

- All the summarized tables mentioned in this SAP will be provided as RTF/HTML or WinWord compatible format (and not as text format). This allows direct insertion of these tables unchanged in the Clinical Study Report.

-Unless otherwise specified, the significance test will be conducted at a two-sided significance level of 0.05

-A third party statistician will independently review the Statistical Analysis Report before disclosing the results.

4.2 Sample size calculation

(ICH E3; 9.7.2. ICH E9; 3.5)

The initial protocol was based in assuming a Pearson correlation coefficient of R=0.5 between pre-baseline and final values of PDSS, two pre-baseline measurements [(V1+V2)/2] and two post-baseline measurements [(V6+ V7)/2]. A standardized mean difference of at least 0.5 (considered as the minimum clinically significant difference) will be detected at a two-sided 0.05 confidence level with a power of at least 0.8, when the sample size/group is at least 20 and 40 patients for control and tested drug groups, respectively.

At a later stage, UNS measuring EDS and cataplexy was used as the primary outcome instead of the PDSS. As UNS was expected with a residual variability larger than PDSS, the initial power of .75 judged as too small, a recalculation of the sample size was performed on a standardized mean difference of 0.5 on the UNS (considered as the minimum clinically significant difference), a ratio 1:2, and a correlation R=0.4. Under these assumptions, and ANCOVA test at 0.05 two-sided confidence level, the probability of a significant effect will be found with a power of 0.85 when the sample size exceeds 36+72 patients for control and tested drug groups, respectively, thus a total of 108 patients.

4.3 Analysis populations and subgroups

- **Screened Set (ScS)**:

All patients with an Informed consent form signed.

- **Randomised Set (RS)**:

All randomized patients.

- **Full Analysis Set (FAS):**

In accordance with the intention-to-treat principle and Section 5.2.1 of the ICH E9 guideline, all randomised patients who received at least one dose of IMP and who have at least one baseline value of UNS (primary efficacy criterion).

- **Per Protocol Set (PPS):**

Consists of all patients in the FAS without any major protocol deviation and having one total score UNS value at V6 or V7.
A protocol deviation is major when it has a possible effect on the outcome. As to whether a deviation is major or minor is fixed during the Blind Review Meeting.
The Per protocol analysis will be conducted at the condition that its sample size should be less than 90% of the FAS.

- **Safety Set (SAF):**

All patients who receiving at least one treatment dose.

4 subgroups will be used in some analyses, according to:
- Age at baseline ([6-11], [12-18[)
- Sex
- BMI at baseline (underweight , normal, overweight and obese)
- Type of narcolepsy (Type 1 or Type 2)

4.4 Protocol deviations

PDs will be defined as early as possible during the study conduct.
All PDs will be carefully identified through eCRF data before the BDRMs. During the BDRMs held before database lock (DBL), each PD will be classified as leading to exclusion from the PP population (“major”) or not leading to exclusion from the PP population (“minor”).
All PDs related to study inclusion or non-inclusion criteria, conduct of the trial, patient management or patient assessment will be described, such as:
- Violation of inclusion criteria
- Violation of exclusion criteria
- Intake of non-allowed concomitant medications
- Randomisation error
- Lack of compliance
- Etc.

Additionally, the PDs that could not be identified through eCRF data will also be reviewed during the BDRMs. This listing will be based on the PD forms completed by the study team. Any “major” PD or information needed to conduct the statistical analyses will be extracted from the listing and imported in the database.

4.5 Statistical analyses

Details concerning definitions about derived parameters are provided in section 4.1.5.

4.5.1 Disposition and Baseline characteristics

Disposition of patients and baseline characteristics will be described by randomised treatment group, to assess their comparability, and overall (treatment groups pooled).

4.5.1.1 Patients disposition

Disposition of patients, including reasons for withdrawal, will be summarized during the study by randomised treatment, overall and by center, in the RS and in the FAS.
If necessary a listing of non-included randomised patients will be provided.

4.5.1.2 Protocol deviations

Protocol deviations at inclusion or after inclusion on V2-V8 period, will be described in the RS, by category of important deviations (based on ICH E3 guideline and ICH E3 Q&A).

4.5.1.3 Demographic data and other baseline characteristics

Demographic data and other baseline characteristics such as age, sex, contraception, narcolepsy history, ECG, vital signs, Polysomnography and Multiple Sleep Latency Test, Clinical Global Impression of baseline Severity (CGI-S) and baseline value of endpoints characteristics will be described.
As the populations RS and FAS are identical, only description in the RS will be done.
All previous treatments will be described.

4.5.2 Treatments of patients

Details concerning definitions on extent of exposure, treatment compliance and concomitant treatments are provided in section 4.1.5 .

4.5.2.1 Extent of exposure and treatment compliance

Extent of exposure (Study duration (days) and treatment duration (days)) and treatment compliance (%) will be described in the FAS and the SS.
The treatment duration will be described on the period V2-V7.

4.5.2.2 Concomitant treatments

All concomitant treatments (specific and non-specific) taken at inclusion and during the treatment period (over the V2-V7 period) will be described in the RS, by ATC code.
A listing will be provided for every patient documenting the existing concomitant medication with date of start and end of treatment.

4.5.3 Efficacy analysis

All efficacy analyses will be performed by randomised treatment group in the FAS and in the PPS.
For inferential analyses involving several results according to the imputation method (MAR, MNAR and observed cases), if there isn't any missing data in the Per Protocol Set, consequently only results with observed cases will be compute.

4.5.3.1 Primary endpoints

The Ullanlinna Narcolepsy Scale (UNS) is an 11-item scale measuring the intensity and frequency of symptoms of narcolepsy. The first four items address typical manifestations of cataplexy occurring in situation as laughing, feeling glad or angry or any exciting situation. The 7 other items measure the sleep latency and the propensity to fall asleep in various situations. The total severity score varies from 0 to 44.

Definition:
The primary efficacy endpoint is defined as the change from baseline to final value in the total UNS Raw score.

4.5.3.1.1 Primary analysis

Analysis:

In order to meet the primary objective of the study, the superiority of Pitolisant (BF2.649) as compared to placebo on efficacy after a 7-week treatment period will be assessed, from the UNS total score expressed in terms of change from baseline to Final value. After a Multiple Imputation, the main inferential model with the center as random effect described in section 4.1.3.1 will be used. The assumptions underlying the model will be checked.

The difference between treatment groups will be calculated as BF2.649 minus placebo. Thus, considering the main analysis, a negative treatment difference will be in favour of BF2.649.

The ICH-E9(R1) addendum on Estimands and Sensitivity Analysis in Clinical Trials (ICH, 2020) provides a framework for defining an estimand which is characterized by five attributes: population, endpoint, intercurrent events and population-level summary. The primary estimand for the primary endpoint of the study is defined with mixed strategies for intercurrent events and described in this table (based on the estimand terminology).

The following table summarizes the main endpoint according to the Estimand Notation (ICH-E9R1):

Attribute[1]	**Primary definition**		**Rationale (as needed)**
Population	Intended to include narcoleptic children from 6 to less than 18 years of age with/without cataplexy.		
Endpoint	Ullanlinna Narcolepsy Scale (UNS) at V6, V7		
Intercurrent events	**Event**	**Strategy**	**Rationale (as needed)**
	Receipt of assigned study treatment	Treatment policy	Strategy applied to align the estimand with the application of the ITT principle.
	Prohibited concomitant medication during the study treatment period	While on treatment	Strategy applied to assess the treatment effect up to the last observation at or prior to the time of the prohibited concomitant medication, which marks the end of adherence to treatment.
	Premature discontinuation from the study at any time	Hypothetical strategy	Strategy applied to estimate what the treatment effect would have been if patients adhered to the treatment by assuming that missing data arise from a missing at random (MAR) mechanism.
Population summary	Difference in LS means from ANCOVA model between treatment groups.		

Statistical elements:

For this comparison the following elements will be provided in a summary table:

- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.1.2 Sensitivity and other analyses

Sensitivity analyses will be performed to assess the robustness of the primary analysis results:

- Using MNAR option of the MI imputation (removing the treatment effect in the imputation model of the partially-complete data sets).
- Using the same model as for the primary analysis will be performed with the center as fixed effect instead of random effect.
- Using the same model as for the primary analysis will be performed with in addition the oxybate sodium intake and treatment-by- oxybate sodium intake interaction as fixed effects.
- Using an observed Cases analysis: The same model as for the primary analysis will be performed in patients having a value of UNS total score at V6 or V7 (i.e. in this case analysis restricted to completers the double blind period).
- Using the same model as for the primary analysis will be performed by actual treatment group, only if actual group is different from randomized group.

The same statistical elements to estimate the treatment effect as for primary analysis will be provided.
In addition, value and change from baseline descriptive statistics at each visit by treatment group will be provided, overall and by subgroups.

4.5.3.2 Secondary endpoints

4.5.3.2.1 Pediatric Daytime Sleepiness Scale scores (PDSS)

PDSS is a sum score with 8 items leading to a score range of 0-32. A score greater than 13 is considered as abnormal sleepiness.

Analysis:
The difference between placebo and BF2.649 will be studied in the same way as the primary analysis (Including sensitivity analyses).
In addition,

- Descriptive statistics at baseline and at each post-baseline visit by treatment group and an évolution graph of Mean ± SE will be provided.
- Comparison by treatment group of the response to the treatment (between Baseline value and End of treatment value) using a Chi2 test will be provided.

Statistical elements:
For this comparison the following elements will be provided in a summary table:

- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.2.2 Ullanlinna Narcolepsy Scale Cataplexy subscore (UNS-CTP)

This subscore is defined as the sum Sum of the first four items of the UNS. (Cf. 7 Plazzi, G, P.Lehert, 2021)

Analysis:
The difference between placebo and BF2.649 will be studied in the same way as the primary analysis (Including sensitivity analyses), only in the type I Narcolepy patients.
In addition,

- Descriptive statistics at baseline and at each post-baseline visit by treatment group and an évolution graph of Mean ± SE will be provided.

- Descriptive statistics at baseline and at each post-baseline visit by treatment group of the response to the treatment will be provided.

Statistical elements:
For this comparison the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.2.3 Weekly rate of Cataplexy (WRC)

Patients (or assisted by their parents or legal representative or teacher) shall fill in the document on a daily basis throughout the double blind phase (13 weeks) and the 7-day prior to the next visit during the open-label period. At each visit the patient should bring back his/her diary to the site for review by the investigator. The investigator reviewed these records with the patient.
One of calculated parameters is the Weekly rate of cataplexy (partial and total).

Analysis:
A negative Binomial Regression will be used according to the general model (section 2.2) for WRC estimated during the period V6-V7 (Justification Plazzi 2021) only in the type I Narcolepy patients (for other patients we don't have the baseline value). The WRC at baseline will be used as a fixed covariate and center as random (In case of non convergence of the model, this one can be updated).
The same way as the primary analysis (Including sensitivity analyses) will be performed: MAR and MNAR imputation and observed cases.
In addition,
- Descriptive statistics by treatment group during the following intervals: Baseline (D-28 to D0), Final DB last 6 weeks DB4=D21-D63 (i.e. V4-V7), Final DB last 2 weeks: DB2=D49-D63 (i.e. V6-V7) will be provided
- A supportive analysis will be to compare the proportion of patients such that WRC <1 between treatment groups (simple logistic regression providing OR with 95%CI)

Missing data handling:
Multiple Imputation (MI) procedure will be used to impute the missing evaluations (Cf.section 4.1.4.1).

Statistical elements:
For this comparison the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.2.4 Maintenance of Wakefulness Test 30-minute version

The Maintenance of Wakefulness Test (MWT) is used in this study to assess an individual's ability to maintain awake while resisting the pressure to fall asleep. Patients will be administrated four 30-minute MWT sessions at 2-hour intervals at inclusion visit (V2) and at endpoint visit (V7 or the last on-study visit). Each session is terminated either at the first unequivocal onset of sleep defined as above or, if sleep onset is not achieved, after maximum in-bed duration of 30 minutes.

Analysis:
MWT was right censored after 30 minutes, resulting into a truncated asymmetric distribution (Plazzi 2021).
A Cox proportional regression will be conducted in considering time to sleep onset (mean of the 4 tests at final visit) and survival response defined as non-response (0=patient remained awake until 30 minutes) and response (1=patient falling asleep before 30 minutes). Patients early terminating for treatment-related or -unrelated reason will be considered as responders or right-censored at the time of interruption, respectively. Center will be considered as a cluster variable in the Cox model to allow calculation of robust variance for the model assuming clusters with correlated observations [Andersen 1982, Therneau 2000].
This analysis will be done only in patients with at least a baseline value or a final value (V7 or the last on-study visit).
The same way as the primary analysis (Including sensitivity analyses) will be performed: MAR and MNAR imputation, center as fixed effect (in MAR imputation) and observed cases.

Descriptive statistics (mean values for each latency parameter and number of Sleep onset in SP) at baseline and V7 by treatment group will be provided.

Missing data handling:
Multiple Imputation (MI) procedure will be used to impute the missing evaluations (Cf.section 4.1.4.1).

Statistical elements:
For this comparison the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.3 Exploratory analysis

4.5.3.3.1 Ullanlinna Narcolepsy Scale EDS subscore (UNS-EDS)

This sub-score EDS is computed for all patients.

Analysis:
- The difference between BF2.649 and placebo will be studied in the same way as the primary analysis (with only one sensitivity analysis: the same model as for the primary analysis by actual treatment group, only if actual group is different from randomized group).

In addition,
- Descriptive statistics at baseline and at each post-baseline visit by treatment group and an évolution graph of Mean ± SE will be provided.
- Descriptive statistics at baseline and at each post-baseline visit by treatment group of the response to the treatment will be provided.

Missing data handling:
Missing evaluations are not imputated.

Statistical elements:
For this comparison the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between adjusted treatment group means.

- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.3.2 Clinical Global Impression of improvement

Clinical Global Impression of improvement is measured for EDS and cataplexy at V6, V7 and V8.

Definition:
This endpoint is defined as the value at the end of each period (V2-V7 and V7-V8) in Global Improvement (CGI-I).

Analyses:
The difference between placebo and BF2.649 will be studied at end of each period using a two-sided Student's t-test for independent samples and a Mann-Whitney test.

Missing data handling:
Missing evaluations are not imputated.

Statistical elements:
For these comparisons the following elements will be provided in a summary table:
- Estimate (standard error) of the difference between treatment group means
- Two-sided 95% CI of the estimate.
- P-value from two-sided Student's t-test for independent (to be compared to 0.05).
- P-value from Mann-Whitney test (to be compared to 0.05).

4.5.3.3.3 Response to treatment according to the CGI-C

An alternative calculation of CGI-C consists in evaluating the number of patients experiencing at least an improvement (CGI-CI) and defined as responder to therapy.

Definition:
This endpoint is defined as the value at end of each period in the response to treatment (based on CGI-I).

Missing data handling:
Missing evaluations are not imputated.

Analyses:
The difference between placebo and BF2.649 will be studied at end of each period using a Chi-Square test.
In addition, descriptive statistics at baseline and at each post-baseline visit by treatment group.

Statistical elements:
For these comparisons the following elements will be provided in a summary table:
- Estimate (Standard Error) of the difference between the proportions of patients.
- Two-sided 95% CI of the estimate.
- P-value from Chi-2 test (to be compared to 0.05).

4.5.3.3.4 Child and Adolescent Sleepiness Scale (CASS)

ESS version adapted for children (6 and over) and adolescents (CASS) with range [0, 30], A score greater than 14 is considered as abnormal sleepiness. CASS will be evaluated at each study visit of double-blind phase and reviewed by the investigator.

Analysis:
The difference between placebo and BF2.649 will be studied in the same way as the primary analysis (Excluding sensitivity analysis).
In addition,

- Descriptive statistics at baseline and at each post-baseline visit by treatment group will be provided

Missing data handling:
Missing evaluations are not imputated

Statistical elements:
For this comparison the following elements will be provided in a summary table:

- Estimate (standard error) of the difference between adjusted treatment group means.
- Two-sided 95% CI of the estimate.
- p-value (to be compared to 0.05).

4.5.3.3.5 Test of Everyday Attention for Children (TEA-Ch)

The Test of Everyday Attention for Children (TEA-Ch) assesses the cognitive functions with a version A at V1 and a version B at V6.
A reference table « TeaCH_tests» can do the correspondance between the code in databe and the item.

Analysis:
Cumulative percentages will be used in analyses (It's more meaningfull than raw scores to compare children according their sex and age)
The difference between placebo and BF2.649 will be studied, by test, in the same way as the primary analysis (Excluding sensitivity analysis).
In addition,

- Value and change from baseline descriptive statistics by visit, by treatment group will be provided.

Missing data handling:
Missing evaluations are not imputated

4.5.3.3.6 Patient sleep diary

The Weekly rate of cataplexy (partial and total) analysis is already done in the secondary endpoint (section 4.5.3.2.3).
The other calculated parameters are:
- Night duration calculated as the difference between Hour of bedtime and wake-time (NightD)
- Number of diurnal involuntary sleep attacks and episodes of severe daytime sleepiness (NSAE).
- Number of hallucinations and sleep paralysis episodes (NHS).

Analysis:
Descriptive statistics by treatment group during the following intervals: Baseline (D-28 to D0), Final DB last 6 weeks DB4=D21-D63, Final DB last 2 weeks: DB2=D49-D63 will be provided

Missing data handling:
Missing evaluations are not imputated

4.5.3.3.7 Patient's Global Opinion on the effect of treatment

At V7 patients' global opinion will be reported by patient if able to express himself. If not, either their parents or their main teachers should evaluate the Global effect of the treatment by comparing the period prior to that visit with the patient's pre-study condition. The six-level of the PGO will be 3= Marked effect, 2=Moderate, 1=Minimal, 0=No, -1= Minimal Worse, -2=Much worse.
We also use the binary variable PGO-C defined as patient having Improved (when PGO=3, 2 or 1) and considered as responder to therapy

Definition:
This endpoint is defined as the value at V7 in the response to treatment (based on PGO-C).

Analyses:
The difference between placebo and BF2.649 will be studied at V7 using a Chi-2 test.
In addition, descriptive statistics of levels by treatment group will be provided

Missing data handling:
Missing evaluations are not imputated.

Statistical elements:
For these comparisons the following elements will be provided in a summary table:
- Estimate (Standard Error) of the difference between the proportions of patients.
- Two-sided 95% CI of the estimate.
- P-value from Chi-2 test (to be compared to 0.05).

4.5.3.3.8 Withdrawal symptoms questionnaire (DSM IV)

A questionnaire is asked to patients about their feeling of any symptoms during the last week.
Analyses:
Descriptive statistics at each visit (T1, V8 and eventually the premature withdrawal visit) by symptom by treatment group will be provided.

4.5.4 Safety data

All safety analyses will be performed by actual treatment group in the Safety Set.

4.5.4.1 Adverse events

The Adverse events (AEs) were reported with the following information: Description, Visit number when AE was reported, Dose of study drug when AE occurred (10 mg, 20 mg, 40 mg, NA), Date of onset or worsening and end date or ongoing, Intensity (mild, moderate, severe), Frequency (once, intermittent and number and period, continuous, unknown), Action taken with studied drug (none, dose modification, temporary interruption, discontinued), Action taken with event (none, corrective treatment, additional exploration, hospitalization ≤ 24 hr, hospitalization > 24 hr), Serious adverse event (SAE) (yes, no), Outcome (recovered, recovered with sequelae, worsened, not yet recovered, death), Imputability (likely related, possibly related, unlikely related), Etiology (studied drug, concomitant treatment, associated disease, other, unknown)

Number of events, number and percentage of patients reporting at least one event, presented by primary system organ class (SOC) and preferred term (PT), over the period V2-V8, will be provided for:
- AE
- Serious AE,

- Treatment emergent AE (TEAE), Treatment emergent AE (TEAE) during the stable dose period, TEAE leading to IMP withdrawal, TEAE related to IMP, serious TEAE, severe TEAE, non-serious TEAE (EUDRACT), Treatment emergent fatal AE.
- Special Interest TEAE (SITEAE).

TEAE also will be described according to the intensity, dose in the stable dose period and outcome.

In case a same patient presented the same AE (i.e. same AE as defined according to SOC and PT) several times during the study, only one occurrence per patient will be reported, but the total number of events will be given.

In addition, only TEAEs with incidence >5% (i.e. % of patients with >5 % for a given PT in at least one treatment group) will be presented by SOC and PT in the SAF population.

A listing with all the individual AEs with the associated information will be provided.

4.5.4.2 Vital signs

Definition:
The following vital signs and clinical examination will be analysed during the study:

- Weight (kg)
- BMI (kg/m²)
- SBP (mmHg)
- DBP (mmHg)
- heart rate (bpm)

For the description of the values at each planned post-baseline visit, in case of retest, only the first analysable one measured at the visit is taken into account.

Analysis:
They will be described, in terms of value at baseline, value at each post-baseline visit as well as in terms of change from baseline for SBP, DBP and heart rate.
Moreover, BMI will be described by class calculated by the World Health Organization (2007): underweight: < -2SD; normal range: [-2SD; +1 SD]; Overweight:] +1SD; +2SD] and obese: > +2SD, taking into account the last post-baseline value in regard to baseline state.

4.5.4.3 Laboratory evaluations

The laboratory parameters will not be performed by a central laboratory, so quantitative analysis are not possible.

A full laboratory test was performed at screening (V1), at the final double-blind treatment phase evaluation visit (V7) and at early withdrawal visit. Blood samples will be collected and all laboratory tests will be performed according to the Good Laboratory Practice (OECD GLP Principles) by site-dependent Laboratory. The results of laboratory tests will be interpreted by the investigator. Laboratory abnormalities will be defined as laboratory test results that are outside the reference range as defined by the normal range from the testing laboratory. Clinically significant abnormal value at V1 will be determined by the investigator and will exclude patients from study participation. The results of abnormal values of laboratory tests will be reported in the CRF beside the normal range

value given by the testing laboratory, the original document containing all the tested values will be kept in the patient study file.
- Hematology: Red Blood Cells, White Blood Cells (differential count, absolute value), hemoglobin, haematocrit, Mean corpuscular volume (MCV), platelet count.
- Biochemistry: sodium, potassium, chloride, creatinine, alkaline phosphatases, urea, alanine aminotransferase (ALAT), aspartate aminotransferase (ASAT), gamma glutamyl transpeptidase (GGT), total bilirubin, glucose, triglycerides, total cholesterol and prothrombine ratio.

- Laboratory parameters will be analysed by abnormality and by clinically significance of the abnormality (number and percentage of patients in each class)
- A listing of patients with at least one clinically significant abnormality will be provided.

A description in term of status (Positive or Negative) will be computed for the following parameters:
- Only for pubescent female a Serum Pregnancy test (β-hCG) will be described.
- Urinalysis: semi-quantitative ("dipstick") analysis for pH, ketone bodies, proteins, glucose, blood.
- Urine drug testing: test of abuse of drugs will be performed on urine samples from 12 years. Screened drugs are amphetamines, barbiturates, benzodiazepines, cannabinoids (THC) and Opioids.

4.5.4.4 Physical examination

A full physical examination will be performed at screening period (V0 (when applicable) and V1) and at each visit from V2 to V8 (or at early withdrawal visit). Any significant abnormality and change from baseline will be recorded in the CRF and in the source document and will be analyzed as safety parameters. Any significant changes will be reported in the AE pages of the CRF.
A description by visit and by body system of the Physical examination will be provided.

4.5.4.5 Electrocardiogram

ECG will be performed at each visit from the screening visit (V1) to the end of the double blind-phase (V8) or at early withdrawal visit. The main ECG parameters including Fridericia's corrected QTc interval ($QTcF = QT/\sqrt[3]{(60/HR)}$) and intra individual changes in these parameters will be reported and analyzed locally

Definition:
A value is considered as analysable if a non-missing value was obtained from a Sinusal rhythm ECG from the central review (Banook).
For the description of the values at each planned post-baseline visit, in case of retest, only the first analysable one measured at the visit is taken into account.

Analyses:
ECG parameters will be described, in terms of values at baseline, values at each post-baseline visit.
Presence of clinically significant ECG abnormalities (yes/no) will be analysed.
ECG parameters will be described, in terms of values at baseline, values at each post-baseline visit.
Presence of clinically significant ECG abnormalities (yes/no) will be analysed.
Description of the following binary variables (Yes/No) were also presented:

-QTcF450 =Yes if QTcF post-dose > 450msec
-ΔQTcF = Yes if the post-dose versus pre-dose QtcD ≥ 60msec
-QT500 =Yes if the QT post-dose ≥ 500 msec

A data list will document individual values for any patient presenting with an abnormal clinically significant value at any time point post dose. Values will be patient ID, first and second reading, treatment, age, gender, and visit.

4.5.4.6 Childhood depression inventory (CDI)

The Childhood Depression Inventory scale is a self-administrated 27-item questionnaire commonly used in research to assess depression. The short version 2 of 12 items is used in the study. The patient is asked to pick out the statement in each of the 12 items which best describes the way he feels right now at the moment of the assessment. The range of scores for the scale are 0-3 = none or minimal; 4-7 = mild; 8-15 = moderate; ≥ 16 = severe depression.

Analyses:
The CDI will be described, in terms of total score at baseline and at each post-baseline visit.
It also described in term of range of scores.

4.5.4.7 Columbia-suicide severity rating scale (C-SSR)

The Columbia-Suicide Severity Rating Scale (C-SSRS) will be evaluated at all visits of the study. This scale is intended to be used by individuals who have received training in its administration. The questions contained in the C-SSRS are suggested probes. Ultimately, the determination of the presence of suicidal ideation or behavior depends on the judgment of the individual administering the scale. The interview is semi-structured with a flexible format. Questions are provided as helpful tools – it is not required to ask any or all questions – just enough to get the appropriate answer. Most important: gather enough clinical information to determine whether patient is suicidal or not. If it is established that a patient has not engaged in any suicidal behavior and/or ideation, then no further questions are required.

Analyses
The CDI will be described, in terms of Suicidal risk (Yes/No) at baseline and at each post-baseline visit.

4.6 Modifications from the statistical section of the protocol

Not Applicable.

4.7 Interim analysis

Not Applicable.

5 Tables, Listings and Graphs (TLG)

BIOPROJET_P11-06_Mock_UP_TFLs_FinalVersion_V1.0.pdf

6 Template of tables, listings and graphs

BIOPROJET_P11-06_Mock_UP_TFLs_FinalVersion_V1.0.pdf

7 References

- Billiard M et al. EFNS guidelines on management of Narcolepsy. Eur J Neurol 2006;13:1035-48
- Dauvilliers Y et al. Clinical aspects and pathophysiology of narcolepsy. Clinical Neurophysiology 2003;114:2000-2017
- Plazzi, G, P.Lehert, 2021, Comparative evaluation of symptom measurement tools in pediatric Narcolepsy, submitted to Sleep journal.

- Andersen, P. and Gill, R. (1982). Cox's regression model for counting processes, a large sample study. *Annals of Statistics* 10, 1100-1120.
- Therneau, T., Grambsch, P., Modeling Survival Data: Extending the Cox Model. Springer-Verlag, 2000.
- Food and Drug Administration (FDA), Multiple endpoints in clinical trials. Guidance for the industry, , January 2017.
- CPMP Working Party on Efficacy of Medical Products. Biostatistical methodology in clinical trials in applications for marketing authorizations for medical products. Stat Med 1995;14:1659–82.
- ICH E9 Statistical Principles for Clinical Trials, CPMP/ICH/363/96 , Mar 1998